Attorney [redacted]cket No.: 07148-108001 / A21-

APPLICATION

FOR

UNITED STATES LETTERS PATENT

TITLE: FATTY ACID ELONGASE 3-KETOACYL COA SYNTHASE POLYPEPTIDES

APPLICANT: JAN G. JAWORSKI AND BRENDA J. BLACKLOCK

CERTIFICATE OF MAILING BY EXPRESS MAIL

Express Mail Label No. EL 734683886US

I hereby certify under 37 CFR §1.10 that this correspondence is being deposited with the United States Postal Service as Express Mail Post Office to Addressee with sufficient postage on the date indicated below and is addressed to the Commissioner for Patents, Washington, D.C. 20231.

June 8, 2001
Date of Deposit

Signature

Vince Defante
Typed or Printed Name of Person Signing Certificate

Fatty Acid Elongase 3-Ketoacyl CoA Synthase Polypeptides

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims priority under 35 U.S.C. §119(e) of U.S. provisional application serial number 60/210,326, filed June 8, 2000.

TECHNICAL FIELD

This invention relates to enzymes involved in very long chain fatty acid (VLCFA) synthesis, and more particularly to chimeras and mutants of nucleic acid sequences encoding fatty acid elongase 3-ketoacyl CoA synthase polypeptides.

BACKGROUND

Plant seeds accumulate primarily 16- and 18-carbon fatty acids (FA). Plants also synthesize very long chain fatty acids (VLCFA). VLCFAs are saturated or unsaturated monocarboxylic acids with an unbranched even-numbered carbon chain that is greater than 18 carbons in length. Very long chain fatty acids are key components of many biologically important compounds in animals, plants, and microorganisms. For example, in animals, the VLCFA arachidonic acid is a precursor to many prostaglandins. In plants, VLCFAs are major constituents of triacylglycerols in many seed oils, are essential precursors for cuticular wax production, and are utilized in the synthesis of glycosylceramides, a component of the plasma membrane. Important VLCFAs include arachidic acid (C20:0; *i.e.,* a 20 carbon chain with no double bonds), behenic acid (C22:0), erucic acid (C22:1), and lignoceric acid (C24:1).

VLCFAs are not desirable in edible oils. Oilseeds of the *Crucifereae* (*e.g.,* rapeseed) and a few other plants, however, accumulate C20 and C22 fatty acids. Although plant breeders have developed rapeseed lines that have low levels of VLCFAs for edible oil purposes, even lower levels would be desirable. On the other hand, vegetable oils having elevated levels of VLCFAs are desirable for certain industrial uses, including uses as lubricants, fuels and as a feedstock for plastics, pharmaceuticals and cosmetics.

The biosynthesis in plants of saturated fatty acids up to an 18-carbon chain occurs in the chloroplast. C2 units from acyl thioesters are linked sequentially, beginning with the

condensation of acetyl Co-enzyme A (CoA) and malonyl-acyl carrier protein (malonyl-ACP) to form a C4 acyl fatty acid. This condensation reaction is catalyzed by a 3-ketoacyl synthase III (KASIII). The enzyme 3-ketoacyl synthase I (KASI) catalyzes the stepwise condensation of a fatty acyl moiety (C4 to C14) with C2 groups and malonyl-ACP to produce a 3-ketoacyl-ACP product that is 2 carbons longer than the original substrate (C6 to C16). The last condensation reaction in the chloroplast, converting C16 to C18, is catalyzed by 3-ketoacyl synthase II (KASII). 3-ketoacyl moieties are also referred to as ß-ketoacyl moieties.

Each elongation cycle involves three additional enzymatic steps in addition to the condensation reaction discussed above. Briefly, the 3-ketoacyl condensation product is reduced to 3-hydroxyacyl-ACP, dehydrated to the enoyl-ACP, and reduced to an acyl-ACP. The fully reduced fatty acyl-ACP reaction product then serves as the substrate for the next cycle of elongation.

The C18:0 saturated fatty acid (stearic acid) can be desaturated to produce a C18:1 fatty acid (oleic acid), which can be transported out of the chloroplast and converted to a C18:2 fatty acid (linoleic acid) or a C18:3 fatty acid (α-linolenic acid). Stearic acid and oleic acid can also be elongated outside the chloroplast to form VLCFAs. The formation of fatty acids longer than 18 carbons depends on the activity of a fatty acid elongase complex to carry out four reactions similar to those described above for fatty acid synthesis in the chloroplast. The initial reaction is catalyzed by an elongase 3-ketoacyl CoA synthase (elongase KCS) and involves the condensation of a two carbon group from malonyl CoA with a C18:0 or C18:1 fatty acyl CoA substrate. A gene encoding an elongase KCS from *Arabidopsis thaliana* has been identified and designated *FAE1*. See, *e.g.*, U.S. Patent No. 6,124,524. The gene product catalyzes the condensation of oleoyl CoA and malonyl CoA, leading to the conversion of the C18 substrate to a C20:1 product, eicosenoyl CoA. Mutations have been identified in the *A. thaliana FAE1* gene (see WO 96/13582). *A. thaliana* plants carrying a mutation in FAE1 have significant decreases in the levels of VLCFAs in seeds.

SUMMARY

Despite 85% sequence identity at the amino acid level between the *Arabidopsis thaliana* FAE1 polypeptide and the *Brassica napus* polypeptide of GenBank Accession No.

AAB72178, the composition of the oil from *A. thaliana* and *B. napus* seeds suggests that the enzymes may have different substrate specificities and/or catalytic activity. VLCFAs constitute about 22% of the seed oil of *A. thaliana*, whereas VLCFAs constitute about 62% of the seed oil in rape. *A. thaliana* seed oil is primarily eicosenoic acid (about 18%), with a small amount of erucic acid and longer-chain monunusaturated fatty acids (about 2%). In contrast, rapeseed oil has a relatively small amount of eicosenoic acid (about 10%) and relatively larger amounts of erucic acid and longer-chain monunsaturates (about 52%).

The present invention provides novel polypeptides with altered elongase KCS substrate specificity and/or catalytic activity. One such novel polypeptide comprises three polypeptide segments. The amino-terminal first polypeptide segment has membrane-anchoring properties. It is joined to a second polypeptide segment whose amino acid sequence is residues 75-114 of SEQ ID NO:12 or residues 75-114 of SEQ ID NO:14, followed by a third polypeptide segment having at least 40% sequence identity to the C-terminal 392 amino acids of SEQ ID NO:4. Examples of such polypeptides have the amino acid sequences shown in SEQ ID NOS:12 and 14. The third polypeptide segment can have an aspartic acid residue at the position corresponding to amino acid 307 of SEQ ID NO:4. Examples of such polypeptides have the amino acid sequences shown in SEQ ID NOS:20, 22, 34 and 36.

Such polypeptides can catalyze the condensation of a C18 fatty acyl substrate and malonyl CoA, leading to the synthesis of a C20 fatty acyl CoA. The fatty acid substrate can be oleic acid (C18:1), in which case the product formed is eicosenoic acid (C20:1). In some instances, the fatty acid substrate is stearic acid (C18:0) and the product formed therefrom is arachidic acid (C20:0). Such polypeptides often can further catalyze the condensation of malonyl CoA and a C20 fatty acyl substrate, leading to the synthesis of a C22 fatty acyl CoA. The substrate often is eicosenoic acid (C20:1) and the product is erucic acid (C22:1). The ratio of the C22 fatty acid product to the C20 fatty acid product (C22:1/C20:1) resulting from the activity of such polypeptides can be about 0.20 or greater, about 0.30 or greater, about 0.40 or greater, or about 0.50 or greater as measured in a yeast microsome assay.

The invention also features a polypeptide comprising in the amino-terminal to carboxy-terminal direction: a first polypeptide segment that has membrane anchoring properties, joined to a second polypeptide segment that has residues 75-114 of SEQ ID NO:2,

which is in turn joined to a third polypeptide segment that has at least 90% sequence identity to residues 115-506 of SEQ ID NO:4. An example of such a polypeptide has the amino acid sequence of SEQ ID NO:8. Also featured is a polypeptide comprising in the amino-terminal to carboxy-terminal direction: a first polypeptide segment having at least 80% sequence identity to residues 1-74 of SEQ ID NO:2, joined to a second polypeptide segment having residues 76-114 of SEQ ID NO:4, joined to a third polypeptide segment having at least 40% sequence identity to residues 115-506 of SEQ ID NO:4. An example of such a polypeptide has the amino acid sequence of SEQ ID NO:10. In some embodiments of these polypeptides, the third segment has an aspartic acid at the position corresponding to amino acid 307 of said SEQ ID NO:4. Examples of such polypeptides have the amino acid sequences of SEQ ID NO:16 and SEQ ID NO:18.

A plant is also disclosed, comprising at least one exogenous nucleic acid encoding one or more of the novel polypeptides disclosed herein, as well as seeds having such nucleic acids.

Nucleic acid constructs of the invention comprise at least one regulatory element operably linked to the nucleic acid coding sequence for a novel polypeptide. Host cells containing such nucleic acid constructs are disclosed. Such host cells include bacterial cells, fungal cells, insect cells, plant cells and animal cells.

A method of altering very long chain fatty acids in an organism is disclosed. The method comprises introducing an exogenous nucleic acid into the organism. The nucleic acid encodes one or more of the polypeptides described herein. The nucleic acid is expressed in the organism to produce the polypeptide(s), and the very long chain fatty acid content of the organism is increased compared to the very long chain fatty acid content of a corresponding organism that lacks the exogenous nucleic acid or does not express the exogenous nucleic acid. Suitable organisms include fungi (*e.g.*, yeast), plants, animals, insects and bacteria. Such organisms can produce a higher level of erucic acid than a corresponding organism that lacks or does not express the exogenous nucleic acid.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. For example, the one letter and three letter abbreviations for amino acids and the one-letter abbreviations for nucleotides are commonly understood. Although methods and

materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. In addition, the materials, methods and examples are illustrative only and not intended to be limiting. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control.

The details of one or more embodiments of the invention are set forth in the accompanying drawings and the description below. Other features, objects, and advantages of the invention will be apparent from the drawings and detailed description, and from the claims.

BRIEF DESCRIPTION OF SEQUENCES

SEQ ID NO:1 is the nucleotide sequence of the *Arabidopsis thaliana FAEI* gene (GenBank Accession No. U29142).

SEQ ID NO:2 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:1 (GenBank Accession No. AAA70154).

SEQ ID NO:3 is the nucleotide sequence of a *Brassica napus* fatty acid elongase KCS (GenBank Accession No. AF009563).

SEQ ID NO:4 is the amino acid sequence of the *B. napus* polypeptide encoded by SEQ ID NO:3 (GenBank Accession No. AAB72178).

SEQ ID NO:5 is the nucleotide sequence of a *B. napus* fatty acid elongase KCS (GenBank Accession No. U50771).

SEQ ID NO:6 is the amino acid sequence of the *B. napus* polypeptide encoded by SEQ ID NO:5 (GenBank Accession No. AAA96054).

SEQ ID NO:7 is a nucleotide sequence encoding a polypeptide designated At114.

SEQ ID NO:8 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:7.

SEQ ID NO:9 is a nucleotide sequence encoding a polypeptide designated At74.

SEQ ID NO:10 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:9.

SEQ ID NO:11 is a nucleotide sequence encoding a polypeptide designated At114 L91C K92R.

SEQ ID NO:12 is the amino acid sequence of the polypeptide encoded by SEQ ID

NO:11.

SEQ ID NO:13 is a nucleotide sequence encoding a polypeptide designated At114 K92R.

SEQ ID NO:14 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:13.

SEQ ID NO:15 is a nucleotide sequence encoding a polypeptide designated At114 G307D.

SEQ ID NO:16 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:15.

SEQ ID NO:17 is a nucleotide sequence encoding a polypeptide designated At74 G306D.

SEQ ID NO:18 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:17.

SEQ ID NO:19 is a nucleotide sequence encoding a polypeptide designated At114 L91C K92R G307D.

SEQ ID NO:20 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:19.

SEQ ID NO:21 is a nucleotide sequence encoding a polypeptide designated At114 K92R G307D.

SEQ ID NO:22 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:21.

SEQ ID NO:23 is a nucleotide sequence encoding a polypeptide designated At254.

SEQ ID NO:24 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:23.

SEQ ID NO:25 is a nucleotide sequence encoding a polypeptide designated At173.

SEQ ID NO:26 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:25.

SEQ ID NO:27 is a nucleotide sequence encoding a polypeptide designated Bn176.

SEQ ID NO:28 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:27.

SEQ ID NO:29 is a nucleotide sequence encoding a polypeptide designated At399.

SEQ ID NO:30 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:29.

SEQ ID NO:31 is a nucleotide sequence encoding a polypeptide designated Bn399.

SEQ ID NO:32 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:31.

SEQ ID NO:33 is a nucleotide sequence encoding a polypeptide designated Bn G307D.

SEQ ID NO:34 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:33.

SEQ ID NO:35 is a nucleotide sequence encoding a polypeptide designated At K92R.

SEQ ID NO:36 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:35.

SEQ ID NO:37 is a nucleotide sequence encoding a polypeptide designated At254 G307D.

SEQ ID NO:38 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:37.

SEQ ID NO:39 is a nucleotide sequence encoding a polypeptide designated At173 G307D.

SEQ ID NO:40 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:39.

SEQ ID NO:41 is a nucleotide sequence encoding a polypeptide designated Bn399 G307D.

SEQ ID NO:42 is the amino acid sequence of the polypeptide encoded by SEQ ID NO:41.

SEQ ID NO:43 is the 3' chimera-specific primer used in the generation of At173.

SEQ ID NO:44 is the 5' chimera-specific primer used in the generation of At173.

SEQ ID NO:45 is the 3' chimera-specific primer used in the generation of At114.

SEQ ID NO:46 is the 5' chimera-specific primer used in the generation of At114.

SEQ ID NO:47 is the 3' chimera-specific primer used in the generation of At74.

SEQ ID NO:48 is the 5' chimera-specific primer used in the generation of At74.

SEQ ID NO:49 is the 3' chimera-specific primer used in the generation of At114

L91C K92R.

SEQ ID NO:50 is the 5' chimera-specific primer used in the generation of At114 L91C K92R.

SEQ ID NO:51 is the 3' chimera-specific primer used in the generation of At114 K92R.

SEQ ID NO:52 is the 5' chimera-specific primer used in the generation of At114 K92R.

SEQ ID NO:53 is the 5' universal primer used in the generation of At-Bn chimeras.

SEQ ID NO:54 is the 3' universal primer used in the generation of At-Bn chimeras.

SEQ ID NO:55 is the 5' universal primer used in the generation of Bn-At chimeras.

SEQ ID NO:56 is the 3' universal primer used in the generation of Bn-At chimeras.

DESCRIPTION OF DRAWINGS

Figure 1 shows amino acid sequences of *Brassica napus* (Bn) elongase KCS polypeptides, *Arabidopsis thaliana* FAE1 (At) and novel chimeric polypeptides and novel chimeric polypeptides containing site-directed modifications. Sequences corresponding to those derived from At FAEl are underlined. Site-directed modifications are indicated in bold. One of the Bn elongase KCS sequences shown corresponds to GenBank Accession No. AAB72178; the other *B. napus* sequence shown corresponds to a second *B. napus* elongase KCS having GenBank Accession No. AAA96054.

Figure 2 shows nucleotide sequences of Bn elongase KCS, At *FAE1* and novel chimeric nucleic acids and novel chimeric nucleic acids containing site-directed modifications. Sequences corresponding to those derived from At *FAE1* are underlined. Site-directed modifications are indicated in bold. The two Bn elongase KCS nucleic acid sequences shown encode the two Bn polypeptides shown in Fig. 1. The GenBank Accession Numbers are AF009563 and U50771, respectively.

Like reference symbols in the various drawings indicate like elements.

DETAILED DESCRIPTION

Fatty Acid Elongase KCS Polypeptides

In one aspect, the invention provides a polypeptide containing the following segments in the amino-terminal to carboxy-terminal direction: a first polypeptide segment having membrane anchoring properties, joined to a second polypeptide segment having the amino acid sequence of residues of 75-114 of SEQ ID NO:12 or SEQ ID NO:14, joined to a third polypeptide segment having at least 40% sequence identity to the C-terminal approximately 392 amino acids of the *Brassica napus* elongase KCS polypeptide shown in SEQ ID NO:4. For example, polypeptides designated At114 L91C K92R (SEQ ID NO:12) and At114 K92R (SEQ ID NO:14) are provided by the present invention. The primary sequence of the novel polypeptides of the invention are identified by the source and number of amino-terminal residues (*e.g.,* At74 polypeptides have 74 amino-terminal residues from *Arabidopsis thaliana*), and site-directed modifications are indicated by the original amino acid residue, the position of the modification and the new residue (*e.g.,* polypeptides containing a K92R site-directed modification had a K at amino acid position 92 which was modified by site-directed mutagenesis of the nucleic acid to encode an R residue).

The above-described polypeptides include a first polypeptide segment that can serve as a membrane anchor. Such a segment has properties that result in the elongase KCS polypeptide being anchored to a membrane, such as a lipid bilayer, detergent bilayer, micelle, or cell membrane. Possession of membrane anchoring properties may be the result of the primary structure, secondary structure and/or tertiary structure of the segment. For example, the segment may contain one or more transmembrane domain(s). Alternatively, a post-translational modification of an amino acid residue within the segment can result in the polypeptide being anchored to a membrane. Suitable modifications include, but are not limited to, covalent attachment of a lipid (*e.g.,* a glycosyl phosphatidylinositol anchor) or a carbohydrate (*e.g.,* an oligosaccharide). See, Alberts et al., *The Cell*, 2nd Edition, Garland Publishing, New York, pp 284-298 and Lodish et al., *Molecular Cell Biology*, 3rd Edition, Scientific American Books, p. 604 and pp. 688-692. The ability of a segment to serve as a membrane anchor can be demonstrated by observing whether a polypeptide having such a segment co-purifies with a membrane fraction. Alternatively, a segment can be a membrane-anchor if, after fusing it to the second and third segments, it is shown that the polypeptide possesses elongase KCS activity in an *in vitro* yeast microsome assay, since elongase KCS

polypeptides are active when anchored to a membrane. As another alternative, computer algorithms, such as Predict Protein or META Predict Protein (see www.embl-heidelberg.de/predictprotein), can be used to predict the presence of a transmembrane domain within a segment, and hence, the ability of that polypeptide segment to serve as a membrane anchor.

Examples of polypeptide segments that can be membrane anchors include, but are not limited to, amino acids 1-74 of *A. thaliana* FAE1 (SEQ ID NO:2), and amino acid sequences having 40% or greater sequence identity to residues 1-74 of SEQ ID NO:2. For example, amino acids 1-75 of an elongase KCS from *B. napus* (GenBank Accession No. AAB72178), amino acids 1-75 of *B. juncea* protein (EMBL Accession No. CAA71898), amino acids 1-75 of an elongase KCS from *B. napus* (GenBank Accession No. AAA96054), amino acids 29-105 of a putative β-ketoacyl-CoA synthase from *A. thaliana* (GenBank Accession No. AAD22309) and amino acids 8-76 of a fatty acid elongase-like protein from *A. thaliana* (EMBL Accession No. CAB36702) have at least 40% sequence identity to SEQ ID NO:2. In some embodiments, the first polypeptide segment has at least 80% sequence identity, 90% sequence identity, at least 95% sequence identity, or at least 99% sequence identity to amino acids 1-74 of SEQ ID NO:2.

A percent identity for any subject nucleic acid or amino acid sequence (*e.g.,* any of the fatty acid elongase chimeras described herein) relative to another "target" nucleic acid or amino acid sequence can be determined as follows. First, a target nucleic acid or amino acid sequence of the invention can be compared and aligned to a subject nucleic acid or amino acid sequence, preferably using the BLAST 2 Sequences (Bl2seq) program from the stand-alone version of BLASTZ containing BLASTN and BLASTP (*e.g.,* version 2.0.14). The stand-alone version of BLASTZ can be obtained at <www.fr.com> or <www.ncbi.nlm.nih.gov>. Instructions explaining how to use BLASTZ, and specifically the Bl2seq program, can be found in the 'readme' file accompanying BLASTZ. The programs also are described in detail by Karlin et al. (*Proc. Natl. Acad. Sci. USA*, 87:2264 (1990) and 90:5873 (1993)) and Altschul et al. (*Nucl. Acids Res.*, 25:3389 (1997)).

Bl2seq performs a comparison between the subject sequence and a target sequence using either the BLASTN (used to compare nucleic acid sequences) or BLASTP (used to compare amino acid sequences) algorithm. Typically, the default parameters of a

BLOSUM62 scoring matrix, gap existence cost of 11, a per residue cost of 1 and a lambda ratio of 0.85 are used when performing amino acid sequence alignments. The output file contains aligned regions of homology between the target sequence and the subject sequence. Once aligned, a length is determined by counting the number of consecutive nucleotides or amino acid residues (*i.e.,* excluding gaps) from the target sequence that align with sequence from the subject sequence starting with any matched position and ending with any other matched position. A matched position is any position where an identical nucleotide or amino acid residue is present in both the target and subject sequence. Gaps of one or more residues can be inserted into a target or subject sequence to maximize sequence alignments between structurally conserved domains (*e.g.,* α-helices, β-sheets, and loops).

The percent identity over a particular length is determined by counting the number of matched positions over that particular length, dividing that number by the length and multiplying the resulting value by 100. For example, if (*i*) a 1000 nucleotide target sequence is compared to a subject nucleic acid sequence (*e.g.,* SEQ ID NO:21), (*ii*) the Bl2seq program presents 200 nucleotides from the target sequence aligned with a region of the subject sequence where the first and last nucleotides of that 200 nucleotide region are matches, and (*iii*) the number of matches over those 200 aligned nucleotides is 180, then the 1000 nucleotide target sequence contains a length of 200 and a percent identity over that length of 90 (*i.e.,* 180 ÷ 200 x 100 = 90).

It will be appreciated that a nucleic acid or amino acid target sequence that aligns with a subject sequence can result in many different lengths with each length having its own percent identity. It is noted that the percent identity value can be rounded to the nearest tenth. For example, 78.11, 78.12, 78.13, and 78.14 is rounded down to 78.1, while 78.15, 78.16, 78.17, 78.18, and 78.19 is rounded up to 78.2. It is also noted that the length value will always be an integer.

Polypeptides of the invention have a second segment which contains amino acid residues, in particular, the amino acid residue corresponding to position 92 in SEQ ID NO:2, that affect elongase KCS substrate specificity. If the residue at position 92 is an arginine residue, the ratio of the C22:1 product to the C20:1 product is higher than the corresponding ratio observed when the residue is a lysine. Accordingly, the second segment (residues 75-114) of Atl14 L91C K92R and Atl14 K92R both possess an R at position 92. Another example of

such a polypeptide has the amino acid sequence of SEQ ID NO:2, except that the lysine at amino acid residue 92 is replaced with an arginine. This polypeptide, designated At K92R, has the amino acid sequence shown in SEQ ID NO:36.

Some polypeptides of the invention have a third segment that has at least 40% sequence identity to residues 115-506 of SEQ ID NO:4, which are the carboxy-terminal 392 amino acids of the *B. napus* polypeptide. In some embodiments, the third polypeptide segment has at least 50% sequence identity, at least 60% sequence identity, at least 70%, 80%, 90%, 95% or 99% sequence identity to the carboxy-terminal 392 amino acids of SEQ ID NO:4.

In some embodiments, the third segment has an aspartic acid residue at the position corresponding to amino acid residue 307 of SEQ ID NO:4. An aspartic acid residue at this position is useful for increasing the catalytic activity of an elongase KCS, compared to the catalytic activity of an otherwise similar polypeptide that has a glycine at this position. For example, polypeptides designated At114 G307D, At74 G306D, At114 L91C K92R G307D, At114 K92R G307D, At254 G307D, At173 G307D, Bn G307D and Bn399 G307D have an aspartic acid residue at the position corresponding to residue 307 of SEQ ID NO:4. These polypeptides have SEQ ID NOS: 16, 18, 20, 22, 38, 40, 34 and 42, respectively.

In some embodiments, the third segment contains one or more of the following groups of residues: GNTSSSS at positions corresponding to residues 423-429 of SEQ ID NO:4, HAGG(R/K)A at positions corresponding to residues 391-396 of SEQ ID NO:4, or MGCSAG at positions corresponding to residues 221-226 of SEQ ID NO:4. These groups of residues are among those that are conserved among elongase KCS polypeptides and are thus found in preferred embodiments.

Segments of a polypeptide are joined to one another by covalent bonds, typically peptide bonds. The segments can be joined directly, without any intervening residues between two segments. Alternatively, one segment can be joined indirectly to an adjacent segment by amino acid residues that are situated between the two adjacent segments and are themselves covalently joined to the adjacent segments. In some embodiments, there are one, two or three intervening amino acid residues. In other embodiments, there are four, five, six, seven, eight, nine or ten intervening residues.

A polypeptide of the invention optionally can possess additional amino acid residues

at the amino-terminus or the carboxy-terminus. For example, six His-tag or FLAG™ residues may be linked to a polypeptide at the amino-terminus. See, *e.g.,* U.S. Patent Nos. 4,851,341 and 5,001,912. A reporter polypeptide, such as green fluorescent protein, may be fused to the carboxy-terminus. See, for example, U.S. Patent No. 5,491,084.

With respect to polypeptides, "isolated" refers to a polypeptide that constitutes the major component in a mixture of components, *e.g.,* 30% or more, 40% or more, 50% or more, 60% or more, 70% or more, 80% or more, 90% or more, or 95% or more by weight. Isolated polypeptides typically are obtained by purification from an organism that makes the polypeptide, although chemical synthesis is also feasible. As used herein, "enriched" refers to a polypeptide that constitutes 20-30% (by weight) of a mixture of components. Methods of polypeptide purification include, for example, chromatography or immunoaffinity techniques.

A polypeptide of the invention may be detected by sodium dodecyl sulphate (SDS)-polyacrylamide gel electrophoresis followed by Coomassie Blue-staining or Western blot analysis using monoclonal or polyclonal antibodies that have binding affinity for the polypeptide to be detected.

The presence of a polypeptide of the invention may often be detected by measuring elongase KCS activity. An elongase KCS can catalyze the condensation of a C18 fatty acyl substrate and malonyl CoA, leading to the formation of a C20 fatty acyl product. C18 fatty acids include C18:0 (*e.g.,* stearic acid), C18:1 (*e.g.,* oleic acid), C18:2 (*e.g.,* linoleic acid), and C18:3 (*e.g.,* α-linolenic acid). In some embodiments, an elongase KCS can catalyze the conversion of a C20 fatty acyl substrate to a C22 fatty acyl product. An example of a C20:1 fatty acyl substrate is an eicosenoyl substrate. Such a substrate can be converted to a C22:1 fatty acyl product, *e.g.,* an erucyl product.

Some polypeptides may result in an elongase KCS that does not form reaction product(s) at a desired rate. Such elongases and their genes are useful as controls in analyses of product formation by enzymatically active elongase KCS polypeptides. Such inactive elongase KCS polypeptides and their genes can also be useful in studying the regulation (*e.g.,* transcription, translation, and post-translational events) of genes encoding enzymatically active elongase KCS polypeptides. Such elongase KCS polypeptides can be attached to Sepharose beads and used for affinity purification of fatty acyl substrates from

crude preparations. In addition, such elongase KCS polypeptides and their genes can also be useful to develop reagents for various purposes, *e.g.,* immunological reagents to monitor expression of a elongase KCS polypeptides or nucleic acid probes or primers to monitor inheritance of a elongase KCS gene in a plant breeding program.

Products formed in plants by elongase reactions involving an elongase KCS can be subsequently used to form fatty acyl triacylglycerides (TAGs) during seed development. Alternatively, such products can be further elongated to form cuticular lipids, such as waxes.

In yet another aspect, the invention provides a polypeptide containing the following segments in the amino-terminal to carboxy-terminal direction: a first polypeptide segment having at least 80% sequence identity to the first 74 amino acids of the *A. thaliana FAE1* gene product (SEQ ID NO:2), joined to a second polypeptide segment having amino acids 76-114 of SEQ ID NO:4, joined to a third polypeptide segment having at least 40% sequence identity to the C-terminal 392 amino acids of a *B. napus* elongase KCS (SEQ ID NO:4). An example of such a polypeptide is At74 (SEQ ID NO:10). This polypeptide possesses an R residue at position 92. Another example is At74 G306D (SEQ ID NO:18), which has a D residue at position 306.

Another novel polypeptide disclosed herein contains the following segments in the amino-terminal to carboxy-terminal direction: a first polypeptide segment having membrane anchoring properties, joined to a second polypeptide segment corresponding to amino acids 75-114 of SEQ ID NO:2, joined to a third polypeptide segment having at least 90% sequence identity to the C-terminal 392 amino acids of SEQ ID NO:4. An example of such a polypeptide is At114 (SEQ ID NO:8).

The invention also features the following polypeptide, comprising in the amino-terminal to carboxy-terminal direction: (a) a first polypeptide segment having at least 90% sequence identity to residues 1-254 of SEQ ID NO:2, joined to (b) a second polypeptide segment having the amino acid sequence of residues 255-506 of SEQ ID NO:4. An example of such a polypeptide is designated At254 and the amino acid sequence is shown in Fig. 1 and SEQ ID NO:24.

Another novel polypeptide comprises (a) a first polypeptide segment having at least 85% sequence identity to residues 1-173 of SEQ ID NO:2, joined to (b) a second polypeptide segment having the amino acid sequence of residues 174-506 of SEQ ID NO:4. An example

of such a polypeptide is designated At173 and the amino acid sequence is shown in Fig. 1 and SEQ ID NO:26.

Another novel polypeptide comprises: (a) a first polypeptide segment having at least 90% sequence identity to residues 1-399 of SEQ ID NO:2, joined to (b) a second polypeptide segment having amino acid residues 400-506 of SEQ ID NO:4. An example of such a polypeptide is designated At399 and the amino acid sequence is shown in Fig. 1 and SEQ ID NO:30. Such a polypeptide can exhibit a product ratio and catalytic activity resembling that of wild-type At FAE1.

The invention also features the following polypeptide, comprising in the amino-terminal to carboxy-terminal direction: (a) a first polypeptide segment having amino acid residues 1-176 of SEQ ID NO:4, joined to (b) a second polypeptide segment having at least 95% sequence identity to residues 177-506 of SEQ ID NO:2. An example of such a polypeptide is designated Bn176 and the amino acid sequence is shown in Fig. 1 and SEQ ID NO:28. In yeast microsome assays, the Bn176 polypeptide exhibits detectable elongase KCS catalytic activity and a C21:1/C20:1 product ratio of about 0.51.

The invention also features the following polypeptide, comprising in the amino-terminal to carboxy-terminal direction: (a) a first polypeptide segment having amino acid residues 1-399 of SEQ ID NO:4, joined to (b) a second polypeptide segment having at least 95% sequence identity to residues 400-506 of SEQ ID NO:2. An example of such a polypeptide is designated Bn399 and the amino acid sequence is shown in Fig. 1 and SEQ ID NO:32. In yeast microsome assays, the Bn399 polypeptide exhibits detectable elongase KCS catalytic activity and a C21:1/C20:1 product ratio of about 0.35.

Elongase KCS Nucleic Acids and Constructs

The present invention also includes nucleic acids encoding the above-described polypeptides. As used herein, nucleic acid refers to RNA or DNA, including cDNA, synthetic DNA or genomic DNA. The nucleic acids may be single- or double-stranded, and if single-stranded, may be either the coding or non-coding strand. As used herein with respect to nucleic acids, "isolated" refers to (i) a naturally-occurring nucleic acid encoding part or all of a polypeptide of the invention, but free of sequences, *i.e.,* coding sequences, that normally flank one or both sides of the nucleic acid encoding polypeptide in a genome; (ii) a

nucleic acid incorporated into a vector or into the genomic DNA of an organism such that the resulting molecule is not identical to any naturally-occurring vector or genomic DNA; or (iii) a cDNA, a genomic nucleic acid fragment, a fragment produced by polymerase chain reaction (PCR) or a restriction fragment. Specifically excluded from this definition are nucleic acids present in mixtures of nucleic acid molecules or cells.

Examples of such nucleic acids include those encoding polypeptides designated At114, At74, At114 L91C K92R, At114 K92R, At114 G307D, At74 G306D, At114 L91C K92R G307D, At114 K92R G307D, At254, At173, Bn176, At399, Bn399 and At K92R. These nucleic acids have SEQ ID NOS: 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31 and 35, respectively. It should be appreciated that nucleic acids having a nucleotide sequence other than the specific nucleotide sequences disclosed can still encode a polypeptide having the exemplified amino acid sequence. The degeneracy of the genetic code is well known to the art; *i.e.,* for many amino acids, there is more than one nucleotide triplet that serves as the codon for the amino acid.

Further provided are nucleic acid constructs comprising the above-described nucleic acid coding sequences. Such constructs may be incorporated into a cloning vector. Cloning vectors suitable for use in the present invention are commercially available and used routinely by those of ordinary skill. Nucleic acid constructs of the invention may additionally comprise one or more regulatory elements operably linked to a nucleic acid coding sequence. Such regulatory elements may include promoter sequences, enhancer sequences, response elements or inducible elements that modulate expression of a nucleic acid sequence. As used herein, "operably linked" refers to positioning of a regulatory element in a construct relative to a nucleic acid coding sequence in such a way as to permit or facilitate expression of the encoded polypeptide. The choice of element(s) that may be included depends upon several factors, including, but not limited to, replication efficiency, selectability, inducibility, targeting, the level of expression desired, ease of recovery and the ability of the host to perform post-translational modifications.

The term "host" or "host cell" includes not only prokaryotes, such as *E. coli*, but also eukaryotes, such as fungal, insect, plant and animal cells. Animal cells include, for example, COS cells and HeLa cells. Fungal cells include yeast cells, such as *Saccharomyces cereviseae* cells. A host cell can be transformed or transfected with a DNA molecule (*e.g.*, a

vector) using techniques known to those of ordinary skill in this art, such as calcium phosphate or lithium acetate precipitation, electroporation, lipofection and particle bombardment. Host cells containing a vector of the present invention may be used for such purposes as propagating the vector, producing a nucleic acid (*e.g.*, DNA, RNA, antisense RNA) or expressing a polypeptide or fragments thereof.

A nucleic acid encoding a novel polypeptide of the invention may be obtained using standard molecular biology techniques, for example, molecular cloning, DNA synthesis, and the polymerase chain reaction (PCR). PCR refers to a procedure or technique in which target nucleic acids are amplified. PCR can be used to amplify specific sequences from DNA as well as RNA, including sequences from total genomic DNA or total cellular RNA. Various PCR methods are described, for example, in *PCR Primer: A Laboratory Manual*, Dieffenbach, C. & Dveksler, G., Eds., Cold Spring Harbor Laboratory Press, 1995. Generally, sequence information from the ends of the region of interest or beyond is employed to design oligonucleotide primers that are identical or similar in sequence to opposite strands of the template to be amplified. Various PCR strategies are available by which site-specific nucleotide sequence modifications can be introduced into a template nucleic acid.

Nucleic acids of the present invention may be detected by methods such as ethidium bromide staining of agarose gels, Southern or Northern blot hybridization, PCR or *in situ* hybridizations. Hybridization typically involves Southern or Northern blotting (see, for example, sections 9.37-9.52 of Sambrook et al., 1989, *"Molecular Cloning, A Laboratory Manual"*, 2nd Edition, Cold Spring Harbor Press, Plainview; NY). Probes should hybridize under high stringency conditions to a nucleic acid or the complement thereof. High stringency conditions can include the use of low ionic strength and high temperature washes, for example 0.015 M NaCl/0.0015 M sodium citrate (0.1X SSC), 0.1% sodium dodecyl sulfate (SDS) at 65°C. In addition, denaturing agents, such as formamide, can be employed during high stringency hybridization, *e.g.,* 50% formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, 75 mM sodium citrate at 42°C.

Transgenic Plants

The invention provides a plant containing an exogenous nucleic acid that encodes a polypeptide of the invention, *e.g.,* nucleic acids encoding a polypeptide having an amino acid sequence as shown in SEQ ID NOS:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, or 36.

Accordingly, a method according to the invention comprises introducing a nucleic acid construct into a plant cell and producing a plant (and progeny of such a plant) from the transformed cell. Techniques for introducing exogenous nucleic acids into monocotyledonous and dicotyledonous plants are known in the art, and include, without limitation, *Agrobacterium*-mediated transformation, viral vector-mediated transformation, electroporation and particle gun transformation, *e.g.*, U.S. Patents 5,204,253 and 6,013,863. If cell or tissue cultures are used as the recipient tissue for transformation, plants can be regenerated from transformed cultures by techniques known to those skilled in the art. Transgenic plants may be entered into a breeding program, *e.g.*, to introduce a nucleic acid encoding a polypeptide into other lines, to transfer the nucleic acid to other species or for further selection of other desirable traits. Alternatively, transgenic plants may be propagated vegetatively for those species amenable to such techniques. Progeny includes descendants of a particular plant or plant line. Progeny of an instant plant include seeds formed on F_1, F_2, F_3, and subsequent generation plants, or seeds formed on BC_1, BC_2, BC_3, and subsequent generation plants. Seeds produced by a transgenic plant can be grown and then selfed (or outcrossed and selfed) to obtain seeds homozygous for the nucleic acid encoding a novel polypeptide.

In another aspect, the invention provides a method of altering very long chain fatty acids in an organism. The method involves introducing an exogenous nucleic acid into the organism. The organism may be, for example, a yeast or a plant. A nucleic acid construct of the invention can alter the levels of very long chain fatty acids in plant tissues expressing the novel polypeptide, compared to VLCFA levels in corresponding tissues from a plant that does not contain or does not express the polypeptide. A comparison can be made, for example, between a transgenic plant of a plant line and a plant of the same line that lacks the nucleic acid construct or does not express the nucleic acid construct in that tissue. Plants having an altered VLCFA composition may be identified by techniques known to the skilled artisan, *e.g.,* thin layer chromatographic or gas-liquid chromatographic (GLC) analysis of the appropriate plant tissue. Novel polypeptides can catalyze the conversion of oleic acid (18:1)

to eicosenoic acid (20:1), and the conversion of eicosenoic acid to erucic acid (22:1). In some embodiments, the ratio of erucic acid to eicosenoic acid (22:1/20:1) is greater than or equal to 0.20, as measured in the yeast microsome assay described below.

A suitable group of plants with which to practice the invention include dicots, such as alfalfa, soybean, rapeseed (high erucic and canola), safflower, or sunflower, and monocots, such as corn, wheat, rye, barley, rice, or sorghum. Suitable rapeseed species include *B. napus*, *B. rapa*, *B. juncea*, and *B. hirta*. Additional plant species suitable for use in the present invention include *Sinapsis alba*, *Crambe abyssinica*, *Limnanthes douglasii* and *L. alba*.

Suitable tissues in which to express polynucleotides and/or polypeptides of the invention include, without limitation, seeds, stems and leaves. Seeds expressing a novel coding sequence can be used to extract an oil having elevated levels of eicosenoic acid and/or erucic acid. Leaf tissues in which a novel coding sequence can be expressed include cells and tissues of the epidermis, *e.g.,* cells that are involved in forming trichomes. Also of interest are epidermal cells involved in forming the cuticular layer. The cuticular layer comprises various very long chain fatty acids and VLCFA derivatives such as alkanes, esters, alcohols and aldehydes. Increasing the amount of VLCFAs in epidermal cells and tissues may enhance defense mechanisms and drought tolerance of plants.

The invention will be further described in the following examples, which do not limit the scope of the invention described in the claims.

EXAMPLES

Example 1—Construction and Cloning of Nucleic Acids

Nucleic acids encoding chimeric polypeptides were generated by an overlap polymerase chain reaction (PCR) strategy. Horton et al. (1989), *Gene*, 77:61-68 and see Figure 1 of Ho et al. (1989), *Gene*, 77:51-59. Briefly, a first round of PCR products were generated in separate reactions using *Arabidopsis thaliana FAE1* and *Brassica napus* elongase KCS nucleic acid as template. Nucleic acid sequences of the *A. thaliana* FAE1 and *B. napus* elongase KCS templates are shown in SEQ ID NO:1 and 3, respectively. The portion of each template that was amplified corresponded to the segment to be combined in a

desired chimera. The amino-terminal fragment of a given chimera was amplified using a 5' universal primer (sense) and a 3' chimera-specific primer (anti-sense). The carboxy-terminal fragment of a given chimera was amplified with a 5' chimera-specific primer (sense) and a 3' universal primer (anti-sense). Universal primer sequences are shown in Table 1 and SED ID NOS: 53-56. Chimera-specific primer sequences are shown in Table 2 and SEQ ID NOS:43-52. The 5' and 3' universal primers anneal to the 5' and 3' ends of the template nucleic acid, respectively, and contain *Bam*HI and *Eco*RI restriction sites, respectively, for ease in subcloning into an expression vector. The 5' chimera-specific primers are antisense to the amino-terminal template and the 3' chimera-specific primers are antisense to the carboxy-terminal template. The 5' and 3' chimera-specific primers each contain an internal complementary sequence where a switch occurs from the At to Bn sequence, or alternatively, from Bn to At.

The products produced by the first round of PCR were purified, and a second round of PCR was conducted using a mixture of the products from the first round of PCR as template nucleic acid. The appropriate 5' and 3' universal primers were used to generate the chimeric nucleic acid product in the second round PCR. The amplified product was then digested with *Bam*HI and *Eco*RI, ligated into pYES2 (Invitrogen, Carlsbad, CA) and transformed into *E. coli*. pYES2 is a yeast centromere-containing, episomal plasmid that is stably propagated in both *E. coli* and in yeast. Each nucleic acid was inserted downstream of the GAL1 promoter in pYES2. The GAL1 promoter is induced in yeast when galactose is present in the medium and repressed when glucose is present in the growth medium.

Nucleic acids encoding polypeptides with site-directed alterations in the coding sequence were also prepared by overlap PCR, using 5' and 3' chimera-specific primers in which the internal complementary region contained the desired sequence modification.

TABLE 1

Chimera type			
5' portion	3' portion	5' universal primer	3' universal primer
At	Bn	5'-ggggatccatgacgtccgttaacgttaagctcc-3' (SEQ ID NO:53)	5'-ccgaattcttaggacgacgttttggacac-3' (SEQ ID NO:54)
Bn	At	5'-ggggatccatgacgtccattaacgtaaagctcc-3 (SEQ ID NO:55)	5'-ccgaattcttaggacgaccgttttggacatgagtctt-3' (SEQ ID NO:56)

TABLE 2

Chimera	3' chimera-specific primers	5' chimera-specific primers
At173	5'-gcgctcgaaaatctattcaagaaca-3' (SEQ ID NO:43)	5'-gttcttgaatagattttcgagcgcaccgatgat-3' (SEQ ID NO:44)
At114	5'-cggaacggcacgtgtgatgattcgtcct-3' (SEQ ID NO:45)	5'-aggacggatcatcacacgcgacgttccg-3' (SEQ ID NO:46)
At74	5'-cccaaaccggtttacctcgttga-3' (SEQ ID NO:47)	5'-tcaacgaggtaaaccggattggg-3' (SEQ ID NO:48)
At114 L91C K92R	5'-ccgcattgcagagttagtgtctctaaa-3' (SEQ ID NO:49)	5'-tttagagacactaactctgcaatgcgg-3' (SEQ ID NO:50)
At114 K92R	5'-ccaccgcatctcagagttagtgtctct-3' (SEQ ID NO:51)	5'-agagacactaactctgagatgcggtgg-3' (SEQ ID NO:52)

Due to a degeneracy in the primer used to generate the nucleic acids encoding carboxy-terminal sequences from *B. napus*, the amino acid residue at the fifth to last position from the carboxy-terminus in the polypeptides designated At114, At114 L91C K92R, At114 K92R and At254 is a P and the polypeptide designated At74 is a Q at that position as indicated in Fig. 1. The polypeptides designated At173 and At399 may have a P or a Q at this position and are shown as Q in Fig. 1. A Q is found in the wild-type Bn polypeptide sequence at this position. In addition, due to PCR infidelity in the preparation of the nucleic acid encoding At114, the amino acid residue at position 439 of SEQ ID NO:8 may be an A or a T, with an A being found in the wild-type Bn sequence. In addition, PCR infidelity in the preparation of the nucleic acid encoding At114 L91C K92R resulted in the residue at position 119 being an N. Position 119 in the wild-type Bn amino acid sequence is a D. Based on the data presented below, this residue can be either a D or an N without any apparent effect on activity.

Mutagenesis was confirmed by automated DNA sequencing, and each construct was used to transform *S. cerevisiae* strain InvScI (Invitrogen) using a lithium-acetate procedure (Gietz, R. and Woods, R., in *Molecular Genetics of Yeast: Practical Approaches*, Oxford Press, pp. 121-134 (1994)).

Example 2—Fatty Acid Elongase KCS Activity in Yeast Microsomes

Elongase KCS enzymatic activity was analyzed by preparing microsomes from transformed yeast cells and assaying these microsomes *in vitro* for elongase KCS activity. Transformed yeast cells were grown overnight in YPD media at 30°C with vigorous shaking. Complete minimal uracil dropout media (cm-ura) supplemented with galactose (2% weight/volume in 40 ml) was inoculated to an OD_{600} of 0.002 to 0.01. Cultures were grown at 30°C to an OD_{600} of approximately 1.5 to 2.0. Cells were harvested by centrifugation at 5000 x*g* for 10 min and washed with 10 ml ice cold isolation buffer (IB), which contains 80 mM Hepes-KOH (pH 7.2), 5 mM EGTA, 5 mM EDTA, 10 mM KCl, 320 mM sucrose and 2 mM DTT). Cells were then resuspended in enough IB to fill a 1.7 ml tube containing 700 µl of 0.5 µm glass beads and yeast microsomes were isolated from the cells essentially as described in Tillman, T. & Bell, R., *J. Biol. Chem.* 261:9144-9149 (1986). The microsomal membrane pellet was recovered by centrifugation at 252,000 x*g* for 60 min. Microsomal pellets were resuspended in a minimal volume of IB, and the protein concentration adjusted to 2.5 µg μl^{-1} by addition of IB containing 15% glycerol. Microsomes were frozen on dry ice and stored at -80°C. The protein concentration in microsomes was determined by the Bradford method (Bradford, Anal. Biochem., 72:248-54, 1976).

Elongase KCS activity was measured essentially as described in Hlousek-Radojcic, et al., *Plant J.* 8:803-809 (1995). Briefly, the standard elongation reaction mix contained 80 mM Hepes-KOH (pH 7.2), 20 mM $MgCl_2$, 500 µM NADPH, 100 µM malonyl-CoA, 10 µM CoA-SH and 15 µM [^{14}C]18:1-CoA (50 µCi μmol^{-1}). The reaction was initiated by the addition of yeast microsomes (6 µg protein) and the mixture was incubated at 30°C, in a final reaction volume of 25 µl. Reaction time was 10 min unless indicated otherwise.

Methyl esters of the acyl-CoA elongase products were prepared by incubation with 500 µl 2% H_2SO_4/MeOH at 80°C for 2 h. Extracted methyl esters were separated on reverse phase silica gel TLC plates (Analtech, Newark, DE), quantified by phosphorimaging, and analyzed by ImageQuant software (Molecular Dynamics, Inc., Sunnyvale, CA). The detection limit for each product is about 0.001 nmoles/min/mg microsomal protein, depending on the phosphorimage exposure time.

Example 3—Elongase KCS Substrate Specificity

Table 3 is a summary of elongase activity and product ratios of *B. napus* (Bn) and *A.*

thaliana (At) elongase KCS nucleic acid sequences expressed in yeast and assayed as described in Example 2. Microsomes prepared from galactose-induced yeast expressing the indicated nucleic acid were assayed after 10 min for conversion of labeled oleoyl substrate to eicosenoyl product, erucyl product, and lignoceryl product. For convenience, fatty acyl substrates and products are oftentimes referred to as the acid rather than as the acyl or acyl CoA. The ratio of 22:1 product to 20:1 product is also shown. Experiments were performed on 17 individual yeast transformants for each construct.

TABLE 3[1]

	18:1(±sd)	20:1(±sd)	22:1(±sd)	20:1+22:1(±sd)	22:1/20:1(±sd)
B. napus elongase KCS (SEQ ID NO:4)	45±4	3.3±0.4	1.4±0.5	4.8±0.2	0.43±0.11
A. thaliana FAE1 (SEQ ID NO:2)	29±9	6±0.8	1.2±0.2	7.1±0.9	0.20±0.04

[1]Amounts of oleic acid (18:1), eicosenoic acid (20:1), erucic acid (22:1), and the sum of 20:1 and 22:1, are expressed as nmol/mg microsomal protein; ±sd = standard deviation.

Table 4 shows the ratio of 22:1/20:1 products produced by Bn, At, and various chimeric polypeptides after incubation of the microsomes with the labeled 18:1 substrate for 5, 10 or 20 min. The results shown in Table 4 represent 4 different microsome preparations from a single yeast transformant with each construct and 2-3 assays of each microsomal preparation. The At *FAE1* (SEQ ID NO:2) produces about 5 times more eicosenoic acid than erucic acid. In contrast, the Bn elongase KCS (SEQ ID NO:4) produces about 2-3 times more eicosenoic acid than erucic acid. See also Table 3.

The At254, At173 and At114 polypeptides have a 22:1/20:1 product ratio that is similar to that of wild-type At *FAE1*, whereas the At74 polypeptide has a product ratio that is similar to that of wild-type Bn (Table 4). These results indicate that amino acids affecting product specificity are present between residues 75 and 114 of the wild-type At elongase KCS. The At74 gene product possesses the amino acid sequence of the Bn elongase KCS of SEQ ID NO:4 at positions 75 to 114, indicating that amino acids of the Bn elongase KCS that differ from the at *FAE1* in this region contribute to the difference in C22:1/C20:1 product ratio.

TABLE 4

Time (min)	Polypeptide Assayed[1]						
	Bn (SEQ ID NO:4)	At (SEQ ID NO:2)	At254 (SEQ ID NO:24)	At1173 (SEQ ID NO:26)	At1114 (SEQ ID NO:8)	At74 (SEQ ID NO:10)	At1114 L91C K92R (SEQ ID NO:12)
5	0.35±0.07	0.18±0.04	0.14±0.03	0.11±0.07	0.17±0.04	0.42±0.07	0.22±0.02
10	0.33±0.10	0.13±0.01	0.11±0.03	0.08±0.01	0.15±0.03	0.36±0.06	0.20±0.02
20	0.35±0.13	0.13±0.02	0.12±0.03	0.11±0.05	0.16±0.04	0.29±0.05	0.20±0.04

[1] The data are the C22:1/C20:1 ratio ± standard deviation

Site-directed modifications were made to the At1114 or At74 nucleic acid sequence within the region corresponding to residues 75 to 114 in order to determine which amino acids contributed to the altered product ratio. The modified nucleic acids were made according to the overlap PCR strategy described in Example 1 and the constructs were introduced into yeast. Elongase KCS activity was measured as described in Example 2. The results showed that changing the At1114 amino acid sequence from alanine to serine and glutamine to lysine at positions 157 and 163, respectively, resulted in undetectable elongase activity. Likewise, changing serine and isoleucine at positions 93 and 95 within At74 to valine in both positions also resulted in undetectable elongase activity.

However, when the leucine and lysine residues at positions 91 and 92 within the At1114 polypeptide were changed to cysteine and arginine, respectively, the C22:1/C20:1 product ratio of the resulting polypeptide, At1114 L91C K92R, was shifted to more closely resemble that of the wild-type Bn polypeptide (Table 4).

Site-directed modifications were made to the At1114 nucleic acid sequence to generate coding sequences for two new polypeptides, one bearing the leucine to cysteine modification and one bearing the lysine to arginine modification. These polypeptides were designated At1114 L91C and At1114 K92R. The nucleotide sequence of the nucleic acid encoding At1114 K92R is shown in SEQ ID NO:13 and the amino acid sequence of the polypeptide is shown in SEQ ID NO:14. The two nucleic acids were introduced into yeast and the activity of each polypeptide was analyzed in yeast microsome assays. The results showed that the L to C-modified polypeptide, At1114 L91C, had low but detectable catalytic activity. The K to R-modified polypeptide, At1114 K92R, had a higher 22:1/20:1 ratio that approached that of wild-type Bn (Table 5). Results presented are the mean of 1 to 3 individual assays each of

at least 7 separate microsomal preparations.

TABLE 5

	20:1+22:1[1]	(22:1/20:1)
At (SEQ ID NO:2)	16.0 +/- 2.7	0.15 +/- 0.04
Bn (SEQ ID NO:4)	9.8 +/- 3.2	0.32 +/- 0.07
At114 K92R (SEQ ID NO:14)	5.8 +/- 3.1	0.32 +/- 0.09

[1] The sum of the amounts of eicosenoic acid (20:1) and erucic acid (22:1) is expressed as nmole/mg microsomal protein.

Example 4–Elongase KCS Catalytic Activity

Table 6 shows the results of yeast microsome assays of Bn elongase KCS, At *FAE1*, and various chimeric polypeptides for various incubation times. The data in Table 6 show the sum of C20:1 and C22:1 in nmole/mg protein from microsome preparations assayed 2 to 3 times each.

The results indicate that the amount of elongase KCS activity of the wild-type At *FAE1* is about 1.5 to 2 times higher than that of wild-type Bn elongase KCS. The At114 polypeptide has an activity that is intermediate between the wild-type At and wild-type Bn, while the At74 polypeptide has an activity that is lower than that of wild-type Bn enzymes. These results indicate that modifying amino acid residues in the region from position 74 to 114 affects elongase activity.

The activity of the At114 L91C K92R gene product was measured in yeast microsomes and is shown in Table 6. The elongase activity of this polypeptide was higher than that of At114.

TABLE 6[1]

Time (min)	pYES2	Bn (SEQ ID NO:2)	At (SEQ ID NO:4)	At114 (SEQ ID NO:8)	At74 (SEQ ID NO:10)	At114 L91C K92R (SEQ ID NO:12)
0	-2.4±2.4	2.7±1.8	2.4±1.8	1.4±1.0	2.2±0.84	2.3±1.08
5	0.24±0.6	5.1±1.6	8.7±1.2	5.9±0.6	3.2±1.1	6.5±1.0
10	0.78±0.6	7.4±1.8	12.1±0.6	8.8±0.5	4.1±0.8	9.6±1.2
20	0.96±0.6	7.8±2.1	13.7±1.8	10.1±1.2	4.4±0.8	11.5±1.1

45	1.32±0.6	8.1±2.2	14.0±0.6	10.2±1.2	4.6±0.5	12.1±0.9

[1] The data are the sum of the C20:1 and C22:1 elongase products (nmole/mg microsomal protein) ± standard deviation.

The elongase activity of the At114 L91C and At114 K92R polypeptides were also assayed in yeast microsomes. The results indicated that the catalytic activity of the At114 L91C polypeptide was about 15-30% of the activity of At114, whereas the activity of At114 K92R was approximately the same as that of At114.

A yeast microsome assay was carried out to compare the Bn elongase KCS shown in SEQ ID NO:4 and another naturally-occurring elongase KCS from the *B. napus* cultivar Askari. The elongase KCS from Askari has the same sequence as that shown in SEQ ID NO:4, except for a valine at position 4 and an aspartic acid at position 307. The results indicated that the Askari elongase KCS had a higher elongase activity and a higher C22:1/C20:1 ratio that did the Bn elongase KCS of SEQ ID NO:4.

Site-directed modifications to SEQ ID NO:3 were made by the techniques described in Example 2 to generate nucleic acids encoding polypeptides Bn I4V, Bn G307D and Bn I4V G307D. The latter polypeptide has the same amino acid sequence as the naturally occurring Askari elongase KCS. After cloning and transforming of each construct into yeast as described in Example 2, microsome assays were performed. Table 7 presents the results from a single experiment in which elongase activity and product ratios for the elongase KCS constructs were measured. Assays were performed as described in Example 2. The results indicate that changing the residue at position 4 from isoleucine to valine had little or no effect on the elongase activity or the C22:1/C20:1 ratio. On the other hand, the Bn G307D polypeptide had a higher elongase activity and produced more C22:1 product than did the unmodified wild-type Bn polypeptide. The amino acid sequence of Bn G307D is shown in SEQ ID NO:34.

TABLE 7[1]

	18:1	20:1	22:1	24:1	22:1/20:1	20:1+22:1
Bn	47.9	5.5	1.9	0.3	0.35	7.7
Bn I4V	48.4	5.5	2.0	0.4	0.37	7.8
Bn G307D	37.2	6.7	5.4	0.7	0.80	12.7
Bn I4V G307D	41.9	6.5	4.5	0.5	0.68	11.6
B. napus (Ask)	37.6	7.7	6.7	0.8	0.86	15.2

[1] Amounts of oleic acid (18:1), eicosenoic acid (20:1), erucic acid (22:1), lignoceric acid (24:1), and the sum of 20:1 and 22:1, are expressed as nmol/mg microsomal protein.

It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.